

A New Spatial Algorithm Based on Adaptive Fuzzy Neural Network for Prediction of Crustal Motion Velocities in Earthquake Research

Nuray Güneri Tosunoğlu¹ · Ayşen Apaydın²

Received: 30 May 2016/Revised: 1 March 2018/Accepted: 15 March 2018/Published online: 23 March 2018
© Taiwan Fuzzy Systems Association and Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract In earthquake studies, different methods are used in modeling of the crustal motions. In case of obscurity data structure, different approaches are needed in solving motion problems. In this paper, a new spatial algorithm has been developed which is based on adaptive fuzzy neural network (AFNN) approach for the prediction of the crustal motion velocities. In order to find the fuzzy class numbers regarding the network model formed by the fuzzification of the studied area, subtractive clustering algorithm is used. In determining the membership function, utilization of the variogram function which models the relationship that depends on distance among spatial data is proposed. The Marmara Region, Turkey, is used as the case for this study. In order to evaluate the performance of the approach, the kriging method is also utilized in the prediction and the results obtained from both methods are compared based on the mean-square-error criteria. It is observed that the AFNN approach yields results which are as effective as those of kriging. Consequently, it is shown that the AFNN approach will contribute to earthquake studies.

Keywords Earthquake · Crustal motion velocities · Spatial prediction · Variogram · Kriging · Fuzzy logic · Adaptive fuzzy neural network

This paper is based on the Ph.D. thesis of Dr. Nuray Güneri Tosunoğlu [32].

✉ Nuray Güneri Tosunoğlu
nguneri@gazi.edu.tr

¹ Department of International Trade, Gazi University, 06500 Ankara, Turkey

² Department of Insurance and Actuarial Science, Ankara University, 06532 Ankara, Turkey

1 Introduction

Earthquake predicting is one of the difficult issues in the world. The main purpose of the earthquake studies is to predict the probability of occurrence of an earthquake in an area in the most reliable way. There have been many studies which have tried to work out earthquake mechanisms and many others which developed different earthquake parameters. Due to the fact that the stress level which is the most significant parameter of the earthquake cannot be directly measured and no observations can be made inside the earth crust, the earthquake prediction problems are carried out with difficulty [33]. Monitoring the crustal motions and observing the changes in the velocities have a great significance in earthquake prediction. Knowing that the crustal motion velocities result in earthquake gives an idea about when an energy accumulation, which causes earthquake on a fault in a specific region, will probably happen; therefore, it is crucial to follow the crustal motions and to observe the variations in velocities.

Geodesic deformation networks need to be established in order to determine the movements and velocities of the crust. The data obtained from the deformation networks are located on a regional area. Spatial statistics are used in the analysis of this data. One of the methods used for spatial statistics is kriging. Kriging is optimal interpolation based on regression. Measurements made for earthquake predictions contain uncertainties arising from reasons such as environmental impacts, insufficiencies in human senses, malfunctioning of the measuring devices and changes in the structure of the data. Kriging does not consider the uncertainty in measurements. Neural networks and fuzzy systems can be used in earthquake prediction due to their ability to solve the problems related to these uncertainties.

Giacinto et al. [15] used neural networks in order to evaluate the risk of earthquake in regions where the risk is already existent. Muller et al. [24] classified seismic events with low magnitude, which had been recorded in France by seismometer network, as fuzzy. Huang and Leung [17] suggested using fuzzy neural networks to estimate correlation between earthquake field and magnitude. It has also been observed that artificial neural networks were effectively used for defining electrical earthquake signals, for producing response spectrums to artificial earthquakes, for estimating the density of radon in the soil as a signal of earthquake, the magnitude of medium intensity earthquake and the basic motion records of the crust [3, 8, 23, 25, 27]. Bodri [10] evaluated the applicability and benefits of neural networks for earthquake estimation. Fuzzy methods to classify strong ground motion records have proven to be useful [4]. Fuzzy methods are used to predict reservoir-induced earthquake, to keep the record of strong ground motions and to predict the following seismic moment [2, 6, 22, 36]. Baldovino and Dadios [9] showed that the earthquake simulator system, which they had developed by fuzzy logic algorithm, had yielded true, reliable and durable results. Aboonahr et al. [1] used fuzzy logical deduction system in order to define the earthquake potential and seismic zone of Iran-Zagros orogenic belt. Likewise, to analyze the seismic risk in the city of Kunming in China, Andric and Lu [7] suggested a new approach which was dependent on fuzzy logic techniques and probability theory. Ameer et al. [5] used ANFIS to get robust ground motion prediction model.

In this study, we aimed to predict the crustal velocities with the help of AFNN approach. AFNN is a kind of artificial neural network based on Takagi–Sugeno fuzzy inference system. The combination of fuzzy inference systems and neural network learning can help to improve the performance of the earthquake prediction.

In AFNN approach, firstly fuzzy clustering is used for the fuzzification of the studied area. For the fuzzification, membership function is used. Membership function is a function that specifies the degree to which a given input belongs to a set [34]. Depending on the type of membership function, different types of fuzzy sets will be obtained. One of the main difficulties in fuzzy set theory has been with the meaning and measurement of membership functions [11]. In this study, since the data are spatial and their distances to each other are important, it is proposed to select the variogram function as membership function in order to identify this factor in fuzzy logic. The variogram provides a description of the distance-dependent relation between the variables [14].

According to suggested algorithm, fuzzy rules are obtained from the network. Based on these rules, predictions of motion velocities at unobserved points are

calculated. In order to evaluate the performance of the approach, suggested AFNN is compared with kriging in terms of functionality. To make it clearer, the structure of this study is given in Fig. 1.

For the data analysis, spatial prediction techniques are utilized since the data on the velocities of the crustal motions are collected from a spatial region. This technique is presented in Sect. 2 of this study. And the rest of the paper is organized as follows. AFNN Inference System will be explained in Sect. 3. Section 4 discusses algorithm based on AFNN. An application on the prediction of the crustal motion velocities in Marmara Region, Turkey will be presented in Sect. 5. Comparison of the results is given in Sect. 6, and the discussion is given in Sect. 7.

2 Spatial Prediction

The values of spatial variables are evident only in the sampled locations of the study area. When the calculation of the unknown values in the unsampled locations is needed, the known values in the sampled locations are utilized. Calculation of spatial variables in an unsampled location is called prediction [14, 31].

Variogram function is used in spatial prediction. The variogram function, $2\gamma(\mathbf{h})$, is used to characterize the distance-dependent relation between two random variables whose distance between is \mathbf{h} and provided that $Z(\mathbf{x})$ is spatial variable, it is defined as follows:

$$2\gamma(\mathbf{h}) = \text{Var}(Z(\mathbf{x}) - Z(\mathbf{x} + \mathbf{h})) = E[Z(\mathbf{x}) - Z(\mathbf{x} + \mathbf{h})]^2 \tag{1}$$

For the determination of the variogram, first the estimation of semi-variogram obtained from the sampling is calculated as follows:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} (z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h}))^2 \tag{2}$$

In Eq. (2), $N(\mathbf{h})$ shows the number of pairs separated by lag \mathbf{h} . $z(\mathbf{x}_i)$ and $z(\mathbf{x}_i + \mathbf{h})$ are the values for the locations \mathbf{x}_i and $\mathbf{x}_i + \mathbf{h}$, respectively [14, 16, 18, 28, 29]. Obtaining semi-variogram values against each of the \mathbf{h} distances, they

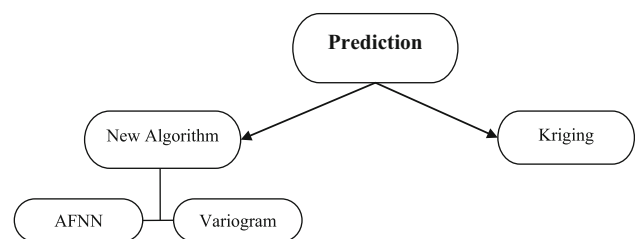


Fig. 1 Structure of the study

will be transferred on the graph and a function adoption is implemented [18, 31]. The most widely used variogram models in the literature are normal, exponential, global, nugget effect, linear, algorithmic, quadratic, proportional quadratic, cubic, power model, wave model and pentagonal models. Variogram function has five defined parameters: C_0 —nugget effect, C —structural variations, $C_0 + C$ —threshold and a —structural distance [31].

Spatial prediction is calculated through the equation below:

$$\hat{z}(\mathbf{x}_0) = \sum_{i=1}^N w_i z(\mathbf{x}_i) \tag{3}$$

In Eq. (3), $\hat{z}(\mathbf{x}_0)$ is the prediction value for the location \mathbf{x}_0 ; $z(\mathbf{x}_i)$ are the values of the variables observed at each \mathbf{x}_i location; w_i are the weight values corresponding to each $z(\mathbf{x}_i)$ and N is the number of points to be used in the prediction of $\hat{z}(\mathbf{x}_0)$ [14, 18]. The weight values given in Eq. (3) were predicted.

Kriging is the determination of weight values in a given prediction in Eq. (3) in a way that the estimation of mean error is zero and the variance is the minimum [31, 35]. After the weights are determined, the prediction value for a given location in the study area is calculated through Eq. (3). In Kriging algorithm, for every new point the weight calculation needs to be repeatedly made [18].

3 AFNN Inference System

Fuzzy inference system is a calculation system based on fuzzy set theory and fuzzy if–then rules. Fuzzy if–then rule is as follows:

$$R_k : \text{If } x \in A_k \text{ then } y \in B_k, \quad k = 1, 2, \dots, K \tag{4}$$

where R_k , k is rule; A_k and B_k are fuzzy sets defined by membership functions; x is linguistic input variable; y is linguistic output variable. The section between “if and then” statements shows the input information (premise), and the section following the then statement shows the output information (consequent) [12, 30]. AFNN system is put forward by Jang [20]. In AFNN system, a relationship is established between the input and output variables in the if–then rule by utilizing learning skills of artificial neural networks and fuzzy rules are determined by means of this relationship. System is a feed-forward network with five layers which are connected to each other by direction links and part of which consists of adaptive neurons. Adaptive neurons have certain parameters. Values of these parameters are determined by means of learning [19, 20]. Fuzzy adaptive network structure with two inputs and two rules is

given in Fig. 2. Operation of network is given below [12, 20]:

Layer 1 Fuzzy sets concerning fuzzy if–then rules are shown by F_1, F_2, F_3 and F_4 . Neurons located in this layer are adaptive, and output value of h neuron is defined as follows, membership function of F_h being μ_{F_h} :

$$\begin{aligned} f_{1,h} &= \mu_{F_h}(x_1), \text{ for } h = 1, 2 \\ f_{1,h} &= \mu_{F_h}(x_2), \text{ for } h = 3, 4 \end{aligned} \tag{5}$$

Layer 2 Neurons located in this layer are demonstrated as Λ_l ($l = 1, \dots, 4$), and they are fixed neurons. Each neuron has two input signals coming from Layer 1. Λ_l is defined as the multiplied of these input signals, and neural functions of this layer are expressed as follows:

$$\begin{aligned} f_{2,1} &= w^1 = \mu_{F_1}(x_1) \cdot \mu_{F_3}(x_2) \\ f_{2,2} &= w^2 = \mu_{F_1}(x_1) \cdot \mu_{F_4}(x_2) \\ f_{2,3} &= w^3 = \mu_{F_2}(x_1) \cdot \mu_{F_3}(x_2) \\ f_{2,4} &= w^4 = \mu_{F_2}(x_1) \cdot \mu_{F_4}(x_2) \end{aligned} \tag{6}$$

Layer 3 Neurons located in this layer are fixed neurons shown by N_l ($l = 1, \dots, 4$). Output value of this layer is the normalization of outputs of Layer 2 and neural function is defined as:

$$f_{3,l} = \bar{w}^l = \frac{w^l}{\sum_{t=1}^4 w^t}, \quad l = 1, \dots, 4 \tag{7}$$

Layer 4 Neurons located in this layer are adaptive neurons of whose neural functions are expressed as follows:

$$f_{4,l} = \bar{w}^l \cdot \hat{Y}^l, \quad l = 1, \dots, 4 \tag{8}$$

\hat{Y}^l is the consequent of a fuzzy if–then rule and is defined as follows:

$$\hat{Y}^l = c_0^l + c_1^l x_1 + c_2^l x_2 \tag{9}$$

c_i^l coefficients in Eq. (9) are fuzzy numbers expressed as $c_i^l = (a_i^l, b_i^l)$ ($i = 0, 1, 2; l = 1, \dots, 4$), and they show consequent parameters.

Layer 5 The single neuron located in this layer is the fixed neuron that calculates the overall output and is calculated as follows:

$$f_{5,1} = f_{\text{output}} = \hat{Y} = \sum_{l=1}^4 \bar{w}^l \cdot \hat{Y}^l = \frac{\sum_{l=1}^4 w_l f_l}{\sum_{l=1}^4 w_l} \tag{10}$$

The aim of AFNN is to achieve the relationship between the input–output data pairs given. This required model is obtained by a learning algorithm. In order to measure the performance of AFNN, different error measures are used. The error measure is defined as the difference between the outputs of the model obtained and the outputs of the target.

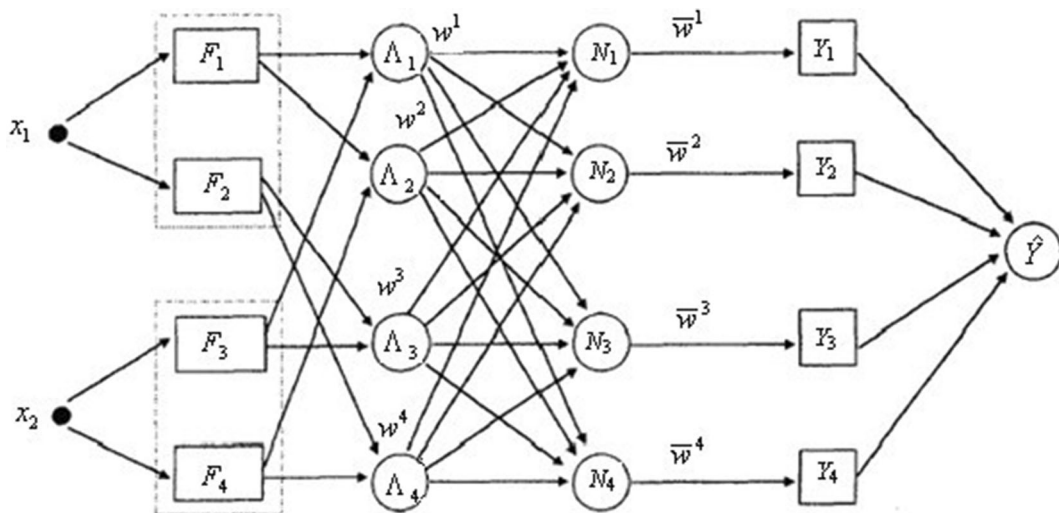


Fig. 2 Structure of fuzzy adaptive network with two inputs and two rules [12]

The training of the network is terminated when this error criterion is less than a prespecified small error.

Different methods are used for the premise and consequent parameters in the training of AFNN. Backpropagation is used for the training of premise parameters, and likelihood linear programming is used for the training of consequent parameters [13, 20, 21].

4 Suggested Algorithm for Spatial Prediction

The main problem in a spatial prediction is to obtain the best prediction of a spatial variable at an unsampled location. To obtain the prediction value for the unsampled location, the observation values of the sampled locations are used. In order to exemplify this problem, a sample prediction problem is given in Fig. 3. The objective here is to obtain the prediction value for the location q with the help of the observation values of the other five locations.

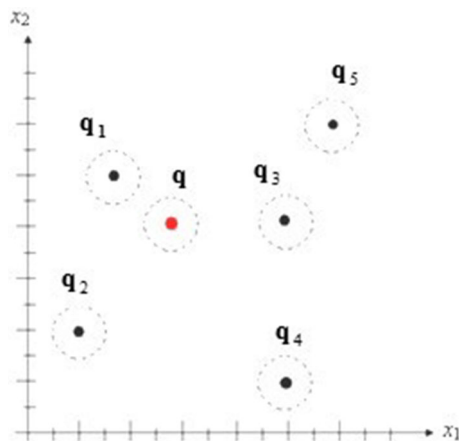


Fig. 3 A sample prediction problem

The value seen in Fig. 3 for the location q_1 covers the area determined around the location q_1 as well. The size of this area, however, is fuzzy. It is thought that this fuzziness should be taken into account in the predictions to be made. But classical spatial statistics methods do not consider the uncertainty. For this reason, AFNN approach is suggested to be used in this prediction problem.

In prediction through AFNN, the independent variables to constitute the inputs of the network are the values of the latitude and the longitude. Spatial prediction with AFNN starts with the determination of sub-cluster numbers of the independent variables and the membership function. Firstly, the study area should be divided into sub-clusters by cluster analysis. Cluster analysis divides data into sub-clusters that are meaningful or useful. There are different methods for cluster analysis. In this study, subtractive clustering algorithm is used. Afterward, the study area should be converted into a fuzzy area. This is known as fuzzification. Fuzzification is the process of changing a real scalar value into a fuzzy value. Membership functions are used in the fuzzification. There are different forms of membership functions such as triangular, trapezoidal, piecewise linear or Gaussian [19]. In the determination of the membership function, a function that models the distance-dependent relation between the variables is suggested. This study is made on spatial variables. For this, variogram function is found to be suitable.

The fuzzy model is constituted based on the Sugeno fuzzy logic method. The fuzzy rules are defined in Table 1.

In Table 1, C_i ($i = 1, \dots, m$) is the sub-clusters for longitude and D_j ($j = 1, \dots, n$) is the sub-clusters for latitude, the x_1 is longitude value, and x_2 is latitude value. Y outputs show the values of spatial variables such as the radon concentration, the intensity of the earthquake and crustal motion velocities. \hat{Y}^l ($l = 1, \dots, n \times m$) values are the

Table 1 Fuzzy rules

Rules	If...then			Model
R^1	If longitude C_1	and	latitude D_1	then $\hat{Y}^1 = c_0^1 + c_1^1 x_1 + c_2^1 x_2$
...
R^n	If longitude C_1	and	latitude D_n	then $\hat{Y}^n = c_0^n + c_1^n x_1 + c_2^n x_2$
R^{n+1}	If longitude C_2	and	latitude D_1	then $\hat{Y}^{n+1} = c_0^{n+1} + c_1^{n+1} x_1 + c_2^{n+1} x_2$
...
$R^{n \times m}$	If longitude C_m	and	latitude D_n	then $\hat{Y}^{n \times m} = c_0^{n \times m} + c_1^{n \times m} x_1 + c_2^{n \times m} x_2$

outputs corresponding to each rule. For each of the $n \times m$ number of rules, the unknown c_i^l ($i = 0, 1, 2, l = 1, \dots, n \times m$) values need to be found.

The algorithm for the determination of most suitable values of the c_i^l coefficients and the premise parameters is defined as follows:

Step 1 With the use of subtractive clustering algorithm, number of fuzzy sub-clusters is obtained.

Step 2 Variogram model suitable for the structure of the data is determined.

Step 3 The variogram model determined in Step 2 will be selected as the membership function.

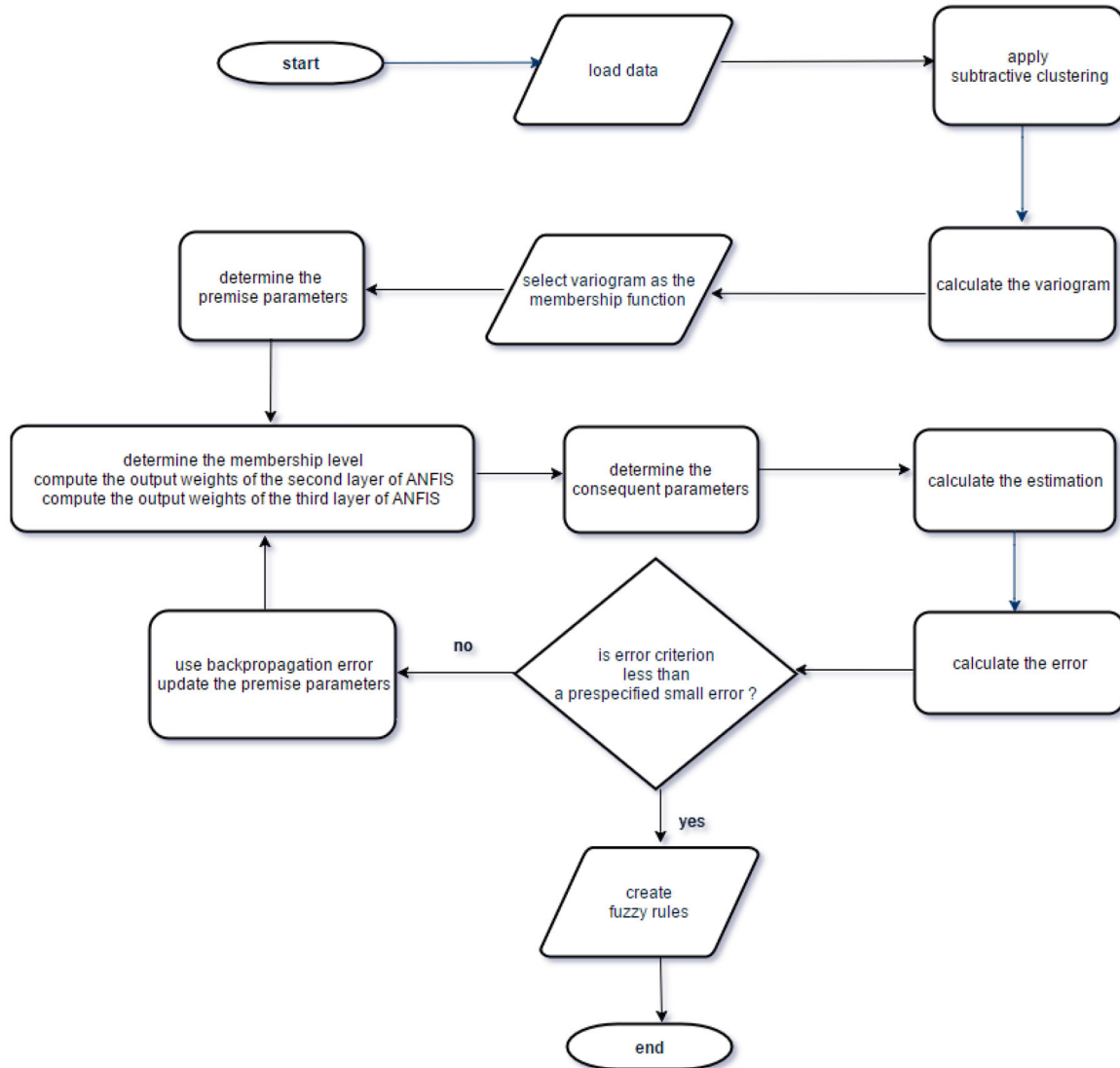


Fig. 4 Flow chart of algorithm

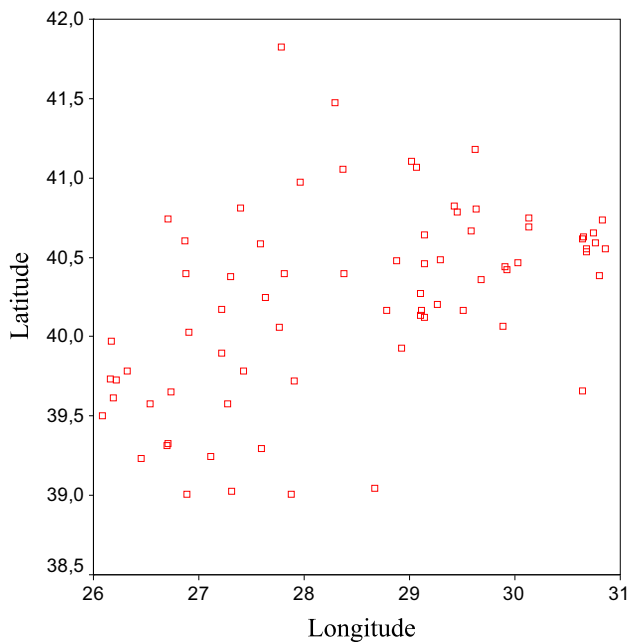


Fig. 5 Locations of the data

Table 2 Descriptive statistics for crustal velocity

Direction	n	Min.	Max.	Mean	SD
North	73	- 13.52	4.92	- 3.86	4.46
East	73	- 24.01	- 19.2	- 16.02	6.84

Step 4 Depending on the number of the fuzzy sub-clusters and the value range of the independent variables determined in the first step, the premise parameters are determined.

Step 5 For every cluster that each of the independent variables belongs to, the value of the membership level is determined. With the use of this membership levels, the weights that are the outputs of the second layer of the adaptive network are obtained. The weights obtained are normalized and the output of the third layer of the adaptive network is obtained.

With the use of the weight values obtained from the third layer, consequent parameter set is determined.

With the use of the consequent parameter set, the models belonging to the fuzzy rules are defined as:

$$\hat{Y}^l = c_0^l + c_1^l x_1 + c_2^l x_2 \tag{11}$$

Using the models constituted and the weights \bar{w}^l determined in Step 5, the estimation values are calculated as follows:

$$\hat{Y} = \sum_{l=1}^m \bar{w}^l \cdot \hat{Y}^l \tag{12}$$

Errors for each observation are calculated and the amount of errors for the model is calculated in:

$$\hat{\varepsilon} = \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2 \tag{13}$$

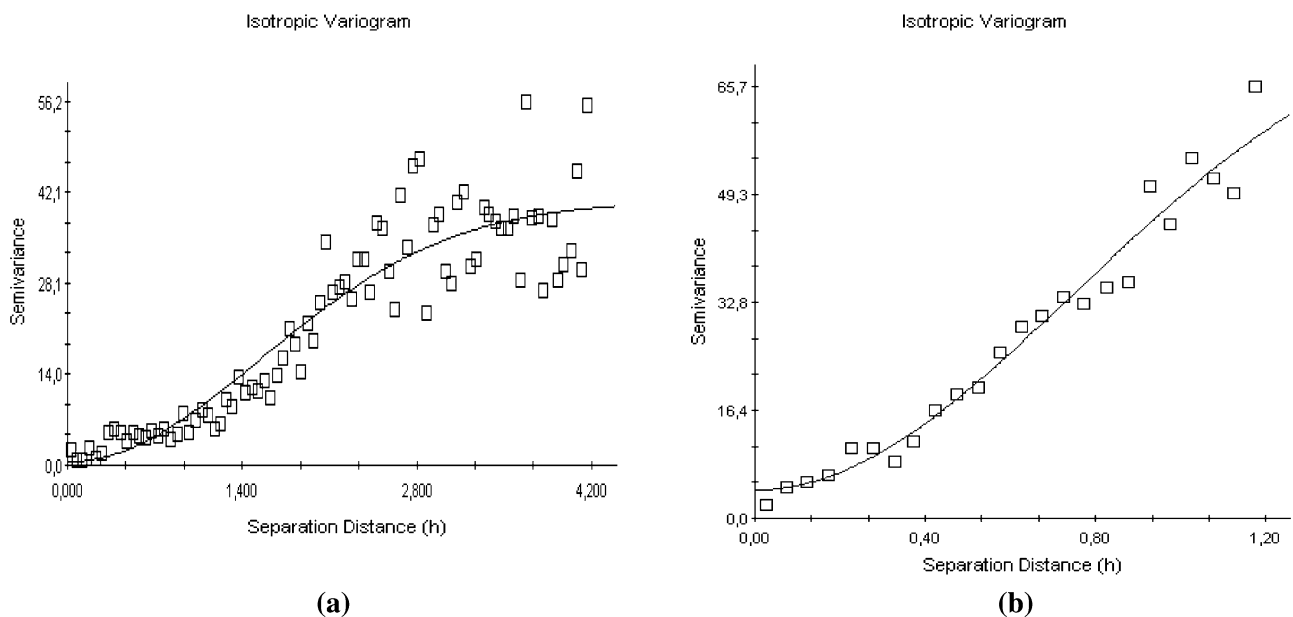


Fig. 6 Sample and model variogram. a North direction and b east direction

Provided ϕ is the amount of error prespecified by the decision maker, if $\hat{\varepsilon} < \phi$, then go to Step 8. If $\hat{\varepsilon} \geq \phi$, then go to Step 6.

Step 6 The backpropagation error that is used in the updating of the premise parameter set is calculated and the premise parameter is updated. The updating formula is

$$\Delta\rho = -\eta \frac{\partial(y_k - \hat{y}_k)^2}{\partial\rho} \quad (14)$$

where ρ is a parameter of the l th neuron at layer r and η is the learning rate.

Table 3 Parameters of variogram model

Direction	C	Parameters	
		C ₀	a
North	40.5	0.3	2.19
East	75.3	4.3	1.05

Table 4 Premise parameters

Premise parameters	Longitude				Latitude			
	Clust1	Clust2	Clust3	Clust4	Clust1	Clust2	Clust3	Clust4
Center	26.08	27.68	29.27	30.86	39.01	39.95	40.89	41.83
Deviation	0.67	0.67	0.67	0.67	0.39	0.39	0.39	0.39

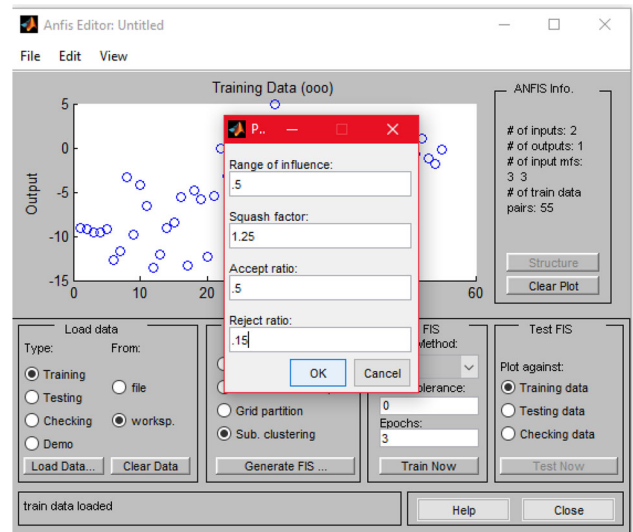
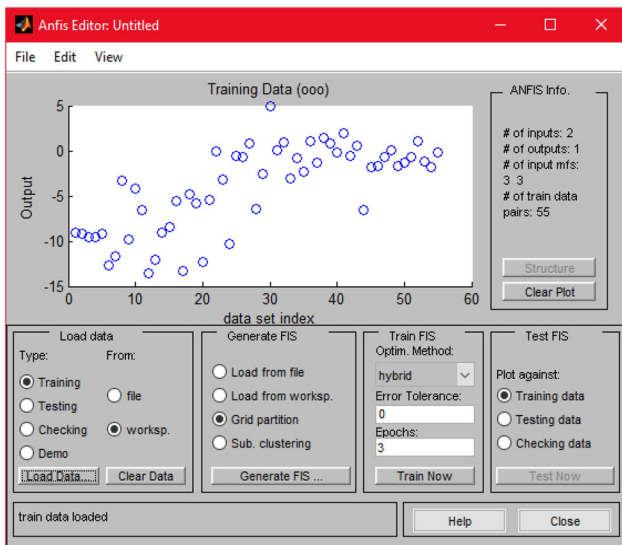
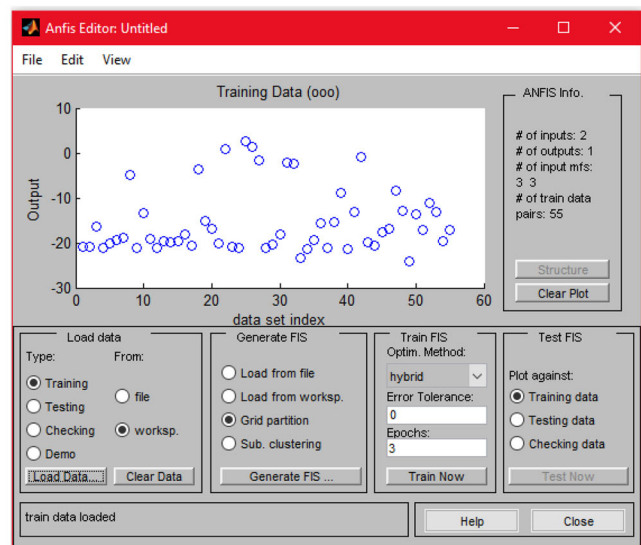


Fig. 8 Subtractive clustering window



(a)



(b)

Fig. 7 Training data. **a** North direction and **b** east direction

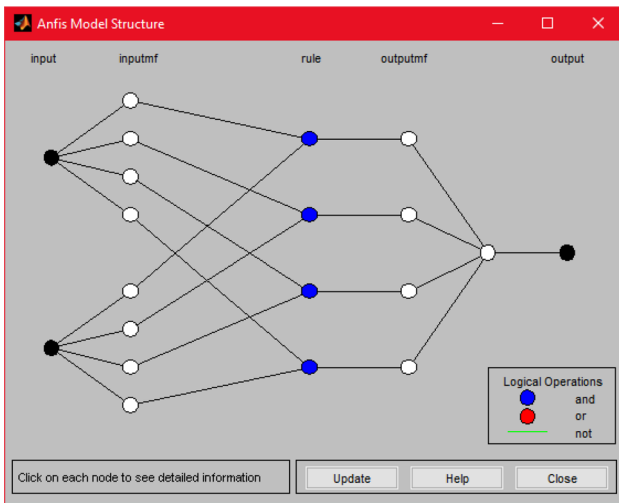


Fig. 9 ANFIS model structure

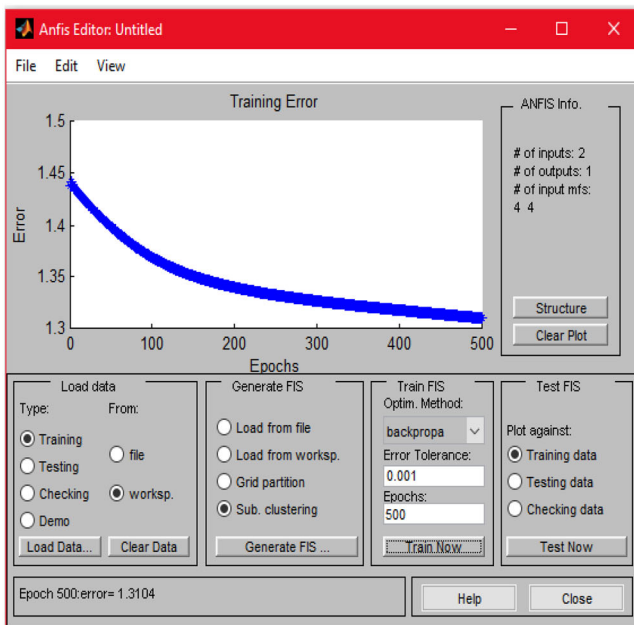
Step 7 Go to Step 4.

Step 8 The algorithm is stopped. The consequent parameter set is set as the parameter for the model to be established. The center and the deviation values, too, correspond to the premise parameter set.

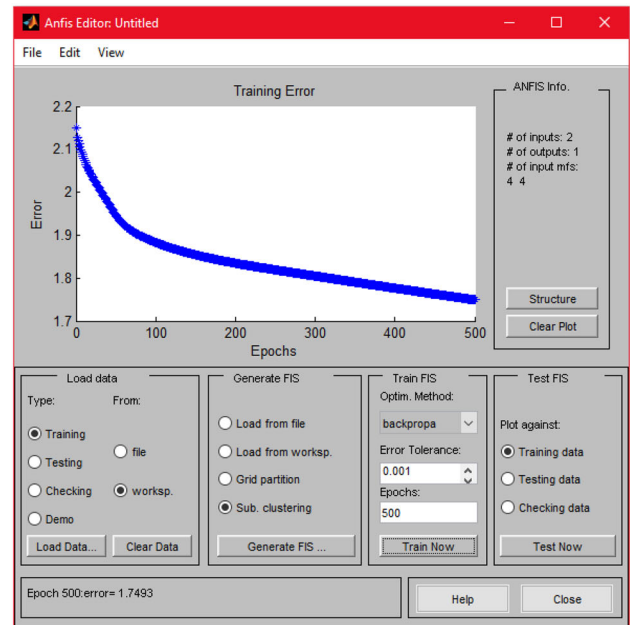
Figure 4 shows the flow chart of algorithm.

5 Application on the Prediction of the Crustal Motion Velocities in Marmara Region, Turkey

For the application, the global positions of the crustal motion velocities and the measurement values corresponding to these positions that are available in the study of Reilinger et al. [26] are used. After the 1999 earthquakes in Turkey, many national and international research projects



(a)



(b)

Fig. 10 Result of training. a North direction. b east direction

Table 5 The center and deviation values for the north direction

	Longitude				Latitude			
	Clust1	Clust2	Clust3	Clust4	Clust1	Clust2	Clust3	Clust4
<i>Center</i>								
Initial	26.08	27.68	29.27	30.86	39.01	39.95	40.89	41.83
Final	26.08	27.42	28.39	30.83	38.95	39.68	40.76	41.35
<i>Deviation</i>								
Initial	0.67	0.67	0.67	0.67	0.39	0.39	0.39	0.39
Final	0.64	0.29	0.32	0.27	0.39	0.27	0.27	0.37

Table 6 The center and deviation values for the east direction

	Longitude				Latitude			
	Clust1	Clust2	Clust3	Clust4	Clust1	Clust2	Clust3	Clust4
<i>Center</i>								
Initial	26.08	27.68	29.27	30.86	39.01	39.95	40.89	41.83
Final	26.00	27.58	29.11	30.98	38.85	40.10	40.98	41.63
<i>Deviation</i>								
Initial	0.67	0.67	0.67	0.67	0.39	0.39	0.39	0.39
Final	0.51	0.56	0.58	0.36	0.09	0.77	0.14	0.55

Table 7 Fuzzy rules for north direction

Rules	If	... and ...	then	Model
R1	If	$x_1 \in C_1$ and $x_2 \in D_1$	then	$\hat{y}_1 = -0.1973 - 0.1958x_1 - 0.0065x_2$
R2	If	$x_1 \in C_1$ and $x_2 \in D_2$	then	$\hat{y}_2 = -0.0546 - 0.1841x_1 - 0.0017x_2$
R3	If	$x_1 \in C_1$ and $x_2 \in D_3$	then	$\hat{y}_3 = -0.1052 - 0.1231x_1 - 0.0038x_2$
R4	If	$x_1 \in C_1$ and $x_2 \in D_4$	then	$\hat{y}_4 = 0.171 + 0.2633x_1 + 0.0064x_2$
R5	If	$x_1 \in C_2$ and $x_2 \in D_1$	then	$\hat{y}_5 = -0.174 - 0.2394x_1 - 0.0062x_2$
R6	If	$x_1 \in C_2$ and $x_2 \in D_2$	then	$\hat{y}_6 = -0.0670 - 0.1637x_1 - 0.0038x_2$
R7	If	$x_1 \in C_2$ and $x_2 \in D_3$	then	$\hat{y}_7 = -0.0346 - 0.1331x_1 - 0.0028x_2$
R8	If	$x_1 \in C_2$ and $x_2 \in D_4$	then	$\hat{y}_8 = -0.0109 - 0.0128x_1 - 0.0006x_2$
R9	If	$x_1 \in C_3$ and $x_2 \in D_1$	then	$\hat{y}_9 = -0.0749 - 0.1141x_1 - 0.0029x_2$
R10	If	$x_1 \in C_3$ and $x_2 \in D_2$	then	$\hat{y}_{10} = 0.0236 - 0.1022x_1 - 0.0030x_2$
R11	If	$x_1 \in C_3$ and $x_2 \in D_3$	then	$\hat{y}_{11} = -0.0886 + 0.1334x_1 + 0.006x_2$
R12	If	$x_1 \in C_3$ and $x_2 \in D_4$	then	$\hat{y}_{12} = -0.117 + 0.0502x_1 + 0.0006x_2$
R13	If	$x_1 \in C_4$ and $x_2 \in D_1$	then	$\hat{y}_{13} = 0.4077 + 0.5231x_1 + 0.0132x_2$
R14	If	$x_1 \in C_4$ and $x_2 \in D_2$	then	$\hat{y}_{14} = -0.0821 - 0.1642x_1 - 0.0035x_2$
R15	If	$x_1 \in C_4$ and $x_2 \in D_3$	then	$\hat{y}_{15} = 0.1319 - 0.14x_1 - 0.0046x_2$
R16	If	$x_1 \in C_4$ and $x_2 \in D_4$	then	$\hat{y}_{16} = 0.1121 + 0.0898x_1 + 0.0023x_2$

have been started in Turkey and it has been aimed to obtain many more parameters about the regional geography. For this reason, the study area is determined to be the region between the latitudes 39°–42° and longitudes 26°–31° which covers the Marmara Region and the surroundings. The 73 values within this region constituted the data set (Fig. 5).

As the motion velocities are given in two directions as northward and eastward, the analyses are made separately for the north and the east directions. Some descriptive statistics about the crustal motion velocities are seen in Table 2.

Table 8 Fuzzy rules for east direction

Rules	If	... and ...	then	Model
R1	If	$x_1 \in C_1$ and $x_2 \in D_1$	then	$\hat{y}_1 = -0.2876 - 0.4228x_1 - 0.0108x_2$
R2	If	$x_1 \in C_1$ and $x_2 \in D_2$	then	$\hat{y}_2 = -0.2854 - 0.3158x_1 - 0.0107x_2$
R3	If	$x_1 \in C_1$ and $x_2 \in D_3$	then	$\hat{y}_3 = 0.196 + 0.3028x_1 + 0.0074x_2$
R4	If	$x_1 \in C_1$ and $x_2 \in D_4$	then	$\hat{y}_4 = 0.3096 + 0.4797x_1 + 0.0118x_2$
R5	If	$x_1 \in C_2$ and $x_2 \in D_1$	then	$\hat{y}_5 = -0.2789 - 0.3943x_1 - 0.0101x_2$
R6	If	$x_1 \in C_2$ and $x_2 \in D_2$	then	$\hat{y}_6 = -0.0107 - 0.494x_1 + 0.004x_2$
R7	If	$x_1 \in C_2$ and $x_2 \in D_3$	then	$\hat{y}_7 = 0.1872 + 0.2807x_1 + 0.0068x_2$
R8	If	$x_1 \in C_2$ and $x_2 \in D_4$	then	$\hat{y}_8 = -0.0244 + 0.0532x_1 - 0.0004x_2$
R9	If	$x_1 \in C_3$ and $x_2 \in D_1$	then	$\hat{y}_9 = -0.2658 - 0.3619x_1 - 0.0093x_2$
R10	If	$x_1 \in C_3$ and $x_2 \in D_2$	then	$\hat{y}_{10} = -0.0561 - 0.0499x_1 - 0.014x_2$
R11	If	$x_1 \in C_3$ and $x_2 \in D_3$	then	$\hat{y}_{11} = -0.0657 - 0.0669x_1 - 0.0021x_2$
R12	If	$x_1 \in C_3$ and $x_2 \in D_4$	then	$\hat{y}_{12} = 0, 1669 + 01931x_1 + 0.0048x_2$
R13	If	$x_1 \in C_4$ and $x_2 \in D_1$	then	$\hat{y}_{13} = -0.0198 - 0.0257x_1 - 0.0006x_2$
R14	If	$x_1 \in C_4$ and $x_2 \in D_2$	then	$\hat{y}_{14} = -0.3416 - 0.3405x_1 - 0.0102x_2$
R15	If	$x_1 \in C_4$ and $x_2 \in D_3$	then	$\hat{y}_{15} = 0.1428 + 0.1879x_1 + 0.0046x_2$
R16	If	$x_1 \in C_4$ and $x_2 \in D_4$	then	$\hat{y}_{16} = 0.3284 + 0.4342x_1 + 0.0107x_2$

5.1 Kriging Application

To compute prediction in kriging method GS+ (Gswin7), Surfer (Version 8.2) is used. First of all, variogram model needs to be determined. Sample variograms are calculated independently from the direction considering the locations of the data. Therefore, average variogram is taken into account in variogram modeling.

The calculated sample variograms and the revised model variograms are shown in Fig. 6a for north direction and in Fig. 6b for east direction.

When the graphs are examined, it is observed that the distribution graph looks like normal distribution for both directions. According to this, for both of the variables, normal distribution model is accepted as the variogram model. The normal distribution model is as follows:

$$\gamma(h) = C_0 + C \left(1 - \exp\left(\frac{-h^2}{a^2}\right) \right) \tag{15}$$

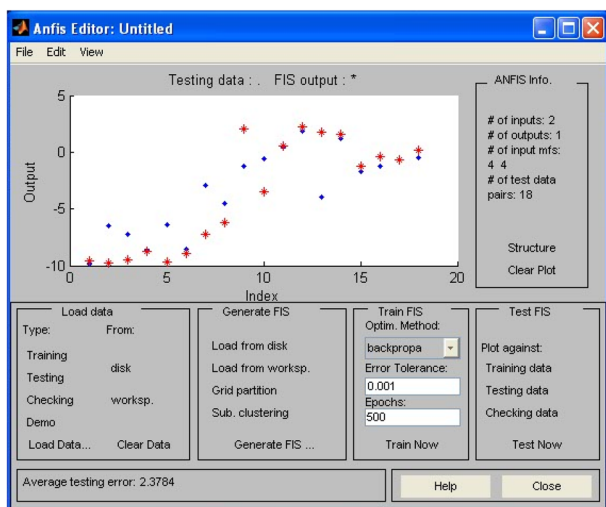
The model parameters are calculated and given in Table 3. After the determination of the variogram model, prediction is computed in kriging method. The predictions are obtained by computing the weights defined in Eq. (3).

The accuracy of the predictions is verified with the use of cross-validation technique. As for the cross-validation, each one of the 73 locations within the data cluster has been taken out from the data cluster in turns and kriging prediction is performed on that location with the help of the other data values. With the calculation of the difference between the prediction values and the measured values, the error values are obtained (Table 4).

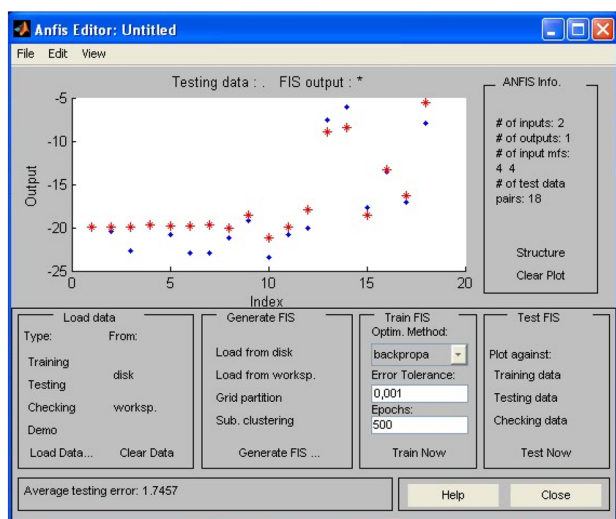
The error value for the kriging prediction is calculated by MSE—mean-square-error criterion, and is found for the north and the east directions, respectively, as follows:

$$MSE_{Kriging-n} = \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N} = 3.4286$$

$$MSE_{Kriging-e} = \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N} = 5.2120$$



(a)



(b)

Fig. 11 Testing data and FIS output. **a** North direction. **b** east direction

5.2 Application of AFNN

For the AFNN application, Anfis Editor under fuzzy logic module of MATLAB software is used. First of all, the data set is separated into two as training set and test set. 75% (55 data) of the data is in the training set and 25% (18 data) is in the test set. The training set is used in the training of the network, and the test set is used in the measuring the performance of the training.

Loaded training data under Anfis Editor are shown in Fig. 7.

Application with the AFNN is shown algorithmically as follows:

Step 1 The numbers of fuzzy sub-clusters for independent latitude and longitude variables are calculated using subtractive clustering algorithm. Parameters required for algorithm are selected as squash factor $\eta = 1.25$, range of influence $r_a = 0.5$, accept ratio $\bar{\epsilon} = 0.5$, reject ratio $\underline{\epsilon} = 0.15$ (Fig. 8).

As a result of the clustering algorithm, the number of sub-clusters for north and east directions is determined as four for both of the latitude and longitude variables. Figure 9 shows the Anfis model structure.

The number of the fuzzy rules to be established as per the numbers of sub-clusters determined is obtained as sixteen through multiplying the numbers of the sub-clusters.

Step 2. For both of the variables, Gaussian distribution model is selected as the variogram model (Eq. 15).

Step 3. According to Step 2, the membership functions are taken as Gaussian distribution and

Step 4. For four cluster, premise parameters are obtained as follows:

Step 5. It is found suitable to perform 500 iterations for training as a result of preliminary tests realized. ϕ is selected as 0.001. Training of the network started. If $\hat{\epsilon} \geq \phi$, then go to Step 6. Otherwise, go to Step 8.

Step 6. The backpropagation error is calculated by using Eq. (14), and the premise parameter set is updated with the training of the network.

Step 7. Go to Step 4.

Step 8. The algorithm is stopped. The result of the training is given in Fig. 10.

The consequent parameter set is set as the parameter for the model. The predictions for the center and the deviation values initially determined and obtained as a result of the training are given in Tables 5 and 6 for north and east directions, respectively.

The fuzzy rules that are formed using the values of consequent parameter set are given in Tables 7 and 8 for north and east directions, respectively. In Tables 7 and 8, the x_1 and x_2 coordinates are the values of the latitude and the longitude, respectively, and C_i, D_j ($i, j = 1, 2, 3, 4$) clusters are the fuzzy sub-clusters for the latitude and longitude values, respectively. The predictions for the crustal motion velocities are calculated by means of these models.

For the evaluation of the results obtained through the method, performance of the test set is considered. Figure 11 shows the testing data and FIS output for test data.

The MSE values for the test set are calculated for the north and the east directions as follows:

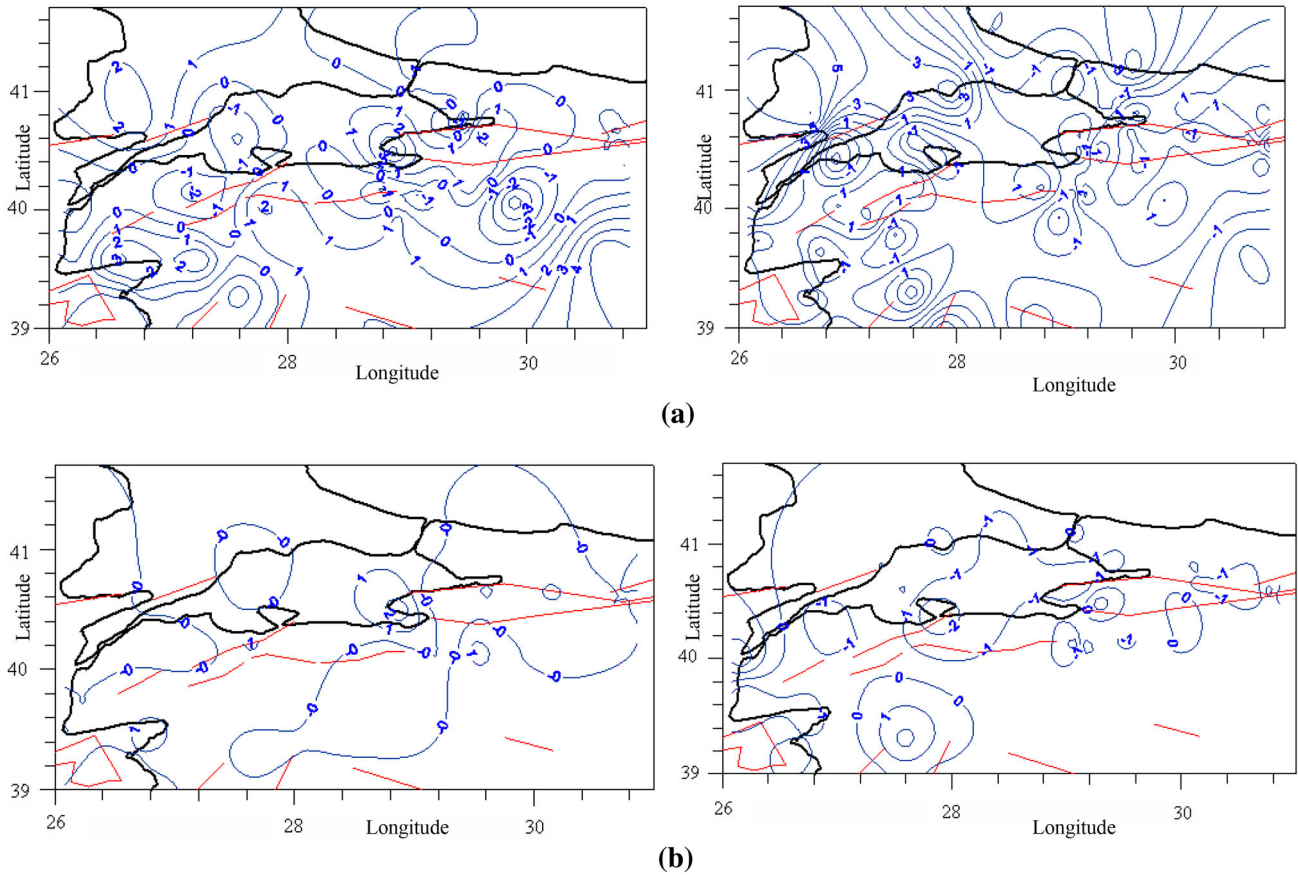


Fig. 12 Errors on crustal velocity prediction (Marmara Region). **a** For Kriging. **b** For AFNN (for north and east directions, respectively) (errors are in mm and equivalent error range is 1 mm)

Table 9 Value of prediction and error of test set for north direction

Latitude	Longitude	Y	\hat{y}_{AFNN}	\hat{e}_{AFNN}	$\hat{y}_{Kriging}$	$\hat{e}_{Kriging}$
26.189	39.614	- 9.85	- 9.57	- 0.28	- 9.32	- 0.53
26.535	39.58	- 6.51	- 9.72	3.21	- 10.14	3.63
26.732	39.653	- 7.26	- 9.48	2.22	- 9.79	2.53
26.91	40.029	- 8.64	- 8.74	0.1	- 7.59	- 1.05
27.269	39.577	- 6.4	- 9.7	3.3	- 9.7	3.3
27.424	39.785	- 8.51	- 8.93	0.42	- 7.61	- 0.9
27.763	40.059	- 2.94	- 7.19	4.25	- 5.27	2.33
27.906	39.722	- 4.53	- 6.21	1.68	- 6.31	1.78
28.373	40.398	- 1.24	2.03	- 3.27	- 0.55	- 0.69
28.923	39.93	- 0.54	- 3.51	2.97	- 1.87	1.33
29.106	40.27	0.44	0.56	- 0.12	- 0.21	0.65
29.146	40.46	1.83	2.22	- 0.39	0.35	1.48
29.451	40.787	- 3.98	1.78	- 5.76	0.95	- 4.93
29.635	40.803	1.23	1.62	- 0.39	- 0.05	1.28
29.929	40.425	- 1.74	- 1.23	- 0.51	- 1.39	- 0.35
30.638	40.614	- 1.21	- 0.4	- 0.81	- 0.78	- 0.43
30.68	40.538	- 0.63	- 0.66	0.03	- 0.91	0.28
30.827	40.735	- 0.52	0.19	- 0.71	- 0.19	- 0.33

$$MSE_{AFNN-n} = \frac{\sum_{k=1}^N (y_k - \hat{y}_k)^2}{N} = 5.6503$$

$$MSE_{AFNN-e} = \frac{\sum_{k=1}^N (y_k - \hat{y}_k)^2}{N} = 3.0725$$

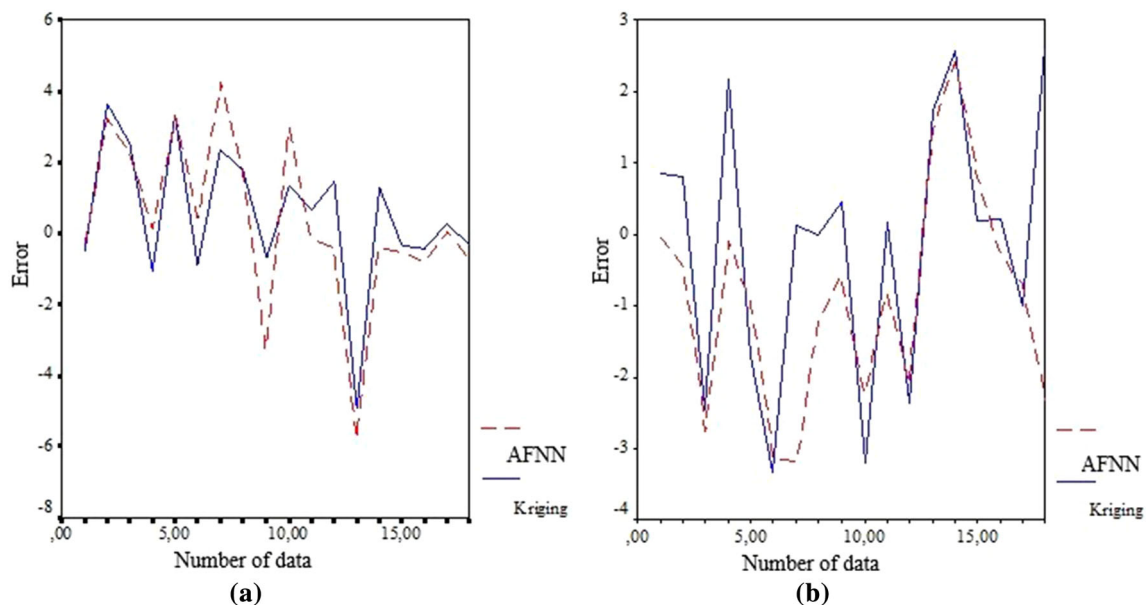
6 Comparison of Results

For the evaluation of the results obtained through both methods, the contour maps are compared. Then the performance of the test set is considered. The errors on the predictions are given in Fig. 12 on the contour map. On the map that is printed on the map of the Marmara Region, the fault lines (red lines) and the errors are shown together. The region is between the latitudes 39°–42° and longitudes 26°–31°.

When the errors are evaluated, it is seen that the contours intensify near the fault lines and larger errors are formed in Kriging method. In AFNN method, however, it is seen that the contours are not very intensified and the errors are small.

Table 10 Value of prediction and error of test set for east direction

Latitude	Longitude	Y	\hat{y}_{AFNN}	\hat{e}_{AFNN}	$\hat{y}_{Kriging}$	$\hat{e}_{Kriging}$
26.189	39.614	-19.94	-19.9	-0.04	-20.8	0.86
26.535	39.58	-20.45	-20	-0.45	-21.26	0.81
26.732	39.653	-22.65	-19.9	-2.75	-20.21	-2.44
26.91	40.029	-19.7	-19.6	-0.1	-21.86	2.16
27.269	39.577	-20.77	-19.8	-0.97	-19.06	-1.71
27.424	39.785	-22.95	-19.8	-3.15	-19.62	-3.33
27.763	40.059	-22.89	-19.7	-3.19	-23.02	0.13
27.906	39.722	-21.2	-20	-1.2	-21.2	0.00
28.373	40.398	-19.18	-18.6	-0.58	-19.63	0.45
28.923	39.93	-23.43	-21.2	-2.23	-20.23	-3.2
29.106	40.27	-20.83	-20	-0.83	-20.99	0.16
29.146	40.46	-20.09	-18	-2.09	-17.72	-2.37
29.451	40.787	-7.5	-8.97	1.47	-9.23	1.73
29.635	40.803	-6.06	-8.47	2.41	-8.62	2.56
29.929	40.425	-17.69	-18.5	0.81	-17.88	0.19
30.638	40.614	-13.58	-13.3	-0.28	-13.79	0.21
30.68	40.538	-17.01	-16.3	-0.71	-16.02	-0.99
30.827	40.735	-7.95	-5.61	-2.34	-10.55	2.6

**Fig. 13** Errors for test data. **a** North direction and **b** east direction

When the results are obtained through both methods, it can be seen that the performance of the test set has significance as well.

The coordinates for the test set and the velocities observed at these coordinates (Y), the prediction values (\hat{y}_{AFNN}) and errors (\hat{e}_{AFNN}) obtained from AFNN approach and the prediction values ($\hat{y}_{Kriging}$) and errors ($\hat{e}_{Kriging}$)

obtained from Kriging method are given in Tables 9 and 10 for the north and east directions, respectively.

From Fig. 13, we can see the errors obtained from both methods for test data.

When the MSE values are compared for the test set, the error obtained in AFNN is found to be larger for the north direction and smaller for the east direction (Table 11).

Table 11 Comparison of MSE for test set

Prediction method	Direction	MSE
AFNN	North	5.6503
	East	3.0725
Kriging	North	4.0356
	East	3.3048

Different from the kriging method, prediction value, in adaptive networks, for a given location can be calculated without the need to make the calculations from scratch with the use of the models belonging to the fuzzy rules obtained.

7 Conclusions

Determination of the crustal motion velocities causing earthquakes is hard and is a prolonged process. Observation stations are established to determine the motion velocities and data are obtained as a result of the measurements made in different periods at these stations. Motion velocities in unknown coordinates can be predicted with the use of motion velocities determined by the observation stations. As there is a spatial relation between the motion velocities, spatial prediction is utilized for this prediction. In this paper, as an alternative to the existing methods used in spatial prediction problems, use of AFNN approach and the variogram function which considers the spatial dependence in selection of the membership function is suggested. In order to determine the efficiency of the system, a comparison is made to the Kriging method. The MSE obtained from the test data showed that the AFNN system had given results as efficient as the kriging method which is known to be the best prediction method. So, it can be said that the ability of AFNN for earthquake prediction is good. When worked with enough data, the AFNN system could be used in solving earthquake prediction problems.

The advantage of the AFNN system over the kriging method is that it has models in hand for predictions. With the help of these models obtained, prediction of the crustal motion velocities in any given location in a study area can be quickly calculated. In kriging method, the weight values used in predictions are calculated based on the distance between the variables. Therefore, the weights need to be recalculated for the prediction of each new spot. It can be said that AFNN has eliminated this problem. When the fuzzy models are obtained with the working of the network, the predictions will be made any time. This is the strength of the AFNN.

Besides the advantages of the AFNN system, there are some points to be careful about during the application

stage. If there are not enough data for the application, models suitable for the AFNN system may not be established. In such circumstances, even though sufficient results for the training set are obtained, the predictions on the sample locations may have big errors. Another point to be watched is the selection of the membership function. In the selection of the membership functions, selection of a variogram model that is not appropriate for the structure of the data will surely affect the performance of the AFNN system, values of the premise parameters and the models for the fuzzy rules to be obtained through the network. For this reason, attention should be given to the selection of the membership function. These are the weaknesses of the AFNN.

Prediction of the crustal motion velocities has significance in terms of earthquake studies. Thanks to the research regarding crustal motions, annual velocity of the deformation in the seismically active portions of the earth's crust can be calculated. The period of time that the deformation reaches the saturation point, taking this velocity into account, can be predicted and earthquakes might be predicted in advance. Therefore, this study is thought to be of contribution to the earthquake prediction studies.

This study is applied in Marmara Region, Turkey. Researchers can similarly use AFNN approach in any spatial region.

References

1. Aboonusr, S.F.G., Zamani, A., Razavipour, F., Boostani, R.: Earthquake hazard assessment in the Zagros Orogenic Belt of Iran using a fuzzy rule-based model. *Acta Geophys.* **65**, 589–605 (2017)
2. Ahumada, A., Altunkaynak, A., Ayoub, A.: Fuzzy logic-based attenuation relationships of strong motion earthquake records. *Expert Syst. Appl.* **42**, 1287–1297 (2015)
3. Alavi, A.H., Gandomi, A.M.: Prediction of principal ground-motion parameters using a hybrid method coupling artificial neural networks and simulated annealing. *Comput. Struct.* **89**, 2176–2194 (2011)
4. Alimoradi, A., Pezeshk, S., Naeim, F., Frigui, H.: Fuzzy pattern classification of strong ground motion records. *J. Earthq. Eng.* **9**(3), 307–332 (2005)
5. Ameer, M., Derras, B., Zendagui, D.: Ground motion prediction model using adaptive neuro-fuzzy inference systems: an example based on the NGA-West 2 data. *Pure Appl. Geophys.* **175**, 1–16 (2017)
6. Andalib, A., Zare, M., Atry, F.: A fuzzy expert system for earthquake prediction, case study: the Zagros range. *Proc. Intell. Transp. Syst.* **15**(3), 1168–1178 (2016)
7. Andrić, J.M., Lu, D.-G.: Fuzzy probabilistic seismic hazard analysis with applications to Kunming city, China. *Nat. Hazards* **89**, 1031–1057 (2017)
8. Asencio-Cortés, G., Martínez-Álvarez, F., Troncoso, A., Morales-Esteban, A.: Medium-large earthquake magnitude prediction

- in Tokyo with artificial neural networks. *Neural Comput. Appl.* **28**, 1043–1055 (2017)
9. Baldovino, R.G., Dadios, E.P.: A hybrid fuzzy logic–PLC-based controller for earthquake simulator system. *JAC III* **20**, 100–105 (2016)
 10. Bodri, B.: A neural-network model for earthquake occurrence. *J. Geodyn.* **32**, 289–310 (2001)
 11. Chen, M.S., Wang, S.W.: Fuzzy clustering analysis for optimizing fuzzy membership functions. *Fuzzy Sets Syst.* **103**, 239–254 (1999)
 12. Cheng, C.B., Lee, E.S.: Applying fuzzy adaptive network to fuzzy regression analysis. *Comput. Math Appl.* **38**, 123–140 (1999)
 13. Cheng, C.B., Lee, E.S.: Fuzzy regression with radial basis function network. *Fuzzy Sets Syst.* **119**, 291–301 (2001)
 14. Cressie, N.A.C.: *Statistics for spatial data*. Wiley, Canada (1993)
 15. Giacinto, G., Paolucci, R., Roli, F.: Application of neural networks and statistical pattern recognition algorithms to earthquake risk evaluation. *Pattern Recogn. Lett.* **18**, 1353–1362 (1997)
 16. Goovaerts, P.: *Geostatistics for natural resources evaluation*. Oxford University Press, New York (1997)
 17. Huang, C., Leung, Y.: Estimating the relationship between isoseismal area and earthquake magnitude by a hybrid fuzzy-neural-network method. *Fuzzy Sets Syst.* **107**, 131–146 (1999)
 18. Isaaks, E.H., Srivastava, R.M.: *An introduction to applied geostatistics*. Oxford University Press, New York (1989)
 19. Ishibuchi, H., Tanaka, H.: Fuzzy neural networks with fuzzy weights and fuzzy biases. In: *Proceedings of 1993 IEEE International Conference on Neural Networks*, San Francisco, pp. 1650–1655 (1993)
 20. Jang, J.-S.R.: ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **23**(3), 665–684 (1993)
 21. Jang, J.-S.R., Sun, C.-T.: Neuro-fuzzy modeling and control. *Proc. IEEE* **83**(3), 378–406 (1995)
 22. Last, M., Rabinowitz, N., Leonard, G.: Predicting the maximum earthquake magnitude from seismic data in Israel and its neighboring countries. *PLoS ONE* **11**(1), e0146 (2016)
 23. Lee, S.C., Han, S.W.: Neural-network-based models for generating artificial earthquakes and response spectra. *Comput. Struct.* **80**, 1627–1638 (2002)
 24. Muller, S., Garda, P., Muller, J.D., Cansi, Y.: Seismic events discrimination by neuro-fuzzy merging of signal and catalogue features. *Phys. Chem. Earth (A)* **24**(3), 201–206 (1999)
 25. Negarestani, A., Setayeshi, S., Maragheh, M.G., Akashe, B.: Estimation of the radon concentration in soil related to the environmental parameters by a modified Adaline neural network. *Appl. Radiat. Isotops* **58**, 269–273 (2003)
 26. Reilinger, R., Mc Clusky, S., Vernant, P., Lawrence, S., Ergintav, S., Cakmak, R., Ozener, H., Kadirov, F., Guliev, I., Stepanyan, R., Nadariya, M., Hahubia, G., Mahmoud, S., Sakr, K., ArRajehi, A., Paradissis, D., Al-Aydrus, A., Prilepin, M., Guseva, T., Evren, E., Dmitrova, A., Filikov, S.V., Gomez, F., Al-Ghazzi, R., Karam, G.: GPS constraints on continental deformation in the Africa–Arabia–Eurasia continental collision zone and implications for the dynamics of plate interactions. *J. Geophys. Res. Solid Earth* **111**(B5): Art. No. B05411 (2006)
 27. Rovithakis, G.A., Vallianatos, F.: A neural network approach to the identification of electric earthquake precursors. *Phys. Chem. Earth (A)* **25**(3), 315–319 (2000)
 28. Sinclair, A.J., Blackwell, G.H.: *Applied Mineral Inventory Estimation*. Cambridge University Press, West Nyack (2002)
 29. Stein, A., Meier, F., Gorte, B.: *Spatial statistics for remote sensing*. Kluwer Academic, Hingham (1999)
 30. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its application to modelling and control. *IEEE Trans. Syst. Man Cybernet.* **15**, 116–132 (1985)
 31. Tercan, A.E., ve Saraç, C.: *Maden Yataklarının Değerlendirilmesinde Jeostatistiksel Yöntemler*, TMMOB Jeoloji Mühendisleri Odası Yayınları: 48, Ankara, Türkiye (1998)
 32. Tosunoğlu, N.G.: Prediction of crustal motion velocities which constitute earthquake by the fuzzy adaptive network in spatial statistics. Ph.D. Thesis, Institute of Science, University of Ankara, Ankara, Turkey (2007) (in Turkish)
 33. Wyss, M.: Why is earthquake prediction research not progressing faster? *Tectonophysics* **338**, 217–223 (2001)
 34. Zadeh, L.A.: Fuzzy sets. *Inf. Control* **8**, 338–353 (1965)
 35. Zhang, J., Yao, N.: The geostatistical framework for spatial prediction. *Geo-Spatial Inf. Sci.* **11**(3), 180–185 (2008)
 36. Zhong, M., Zhang, Q.: Prediction of reservoir-induced earthquake based on fuzzy theory. In: *Proceedings of the Second International Symposium on Networking and Network Security (ISNNS'10)*, Jinggangshan, pp. 101–104 (2010)
- Nuray Güneri Tosunoğlu** received her B.S., M.S. and Ph.D. degree in Statistics from University of Ankara, Turkey, in 1998, 2001 and 2007, respectively. She is currently an associate professor in the Gazi University. Her research interests include artificial intelligence, applied statistics, multi criteria decision making and scientific research methods.
- Ayşen Apaydın** received her B.S., M.S. and Ph.D. degree in Statistics from University of Hacettepe, Turkey, in 1979, 1981 and 1987, respectively. She has been a professor at the Ankara University since 2002. Her research interests include artificial intelligence, optimization, operations research and applied statistics.